

Notas

1. En Cataluña muchas personas pronuncian *RACO* como *racó*, palabra del léxico común que significa rincón. Realmente, un repositorio como éste puede asimilarse a la idea de un rincón donde rebuscar entre mil y un papeles (viejos y nuevos) en busca de un documento concreto o de un hallazgo inesperado.
2. Puede verse el resto de proyectos iniciados por el CBUC (hoy integrado dentro del CSUC) en: <http://www.csuc.cat/ca/biblioteques-cbuc>
3. Datos elaborados a partir de la clasificación de las revistas por materias, disponible en *RACO*. <http://www.raco.cat/index.php/raco/subject>
4. Por ejemplo las *Memorias de la Real Academia de Buenas Letras de Barcelona* se iniciaron en 1868.

"El valor principal de *RACO* es dar acceso abierto, gratuito, universal e ilimitado a la mayoría de los artículos de revista publicados en Cataluña"

Bibliografía

Anglada, Lluís; Reoyo, Sandra (2005). "Els dipòsits electrònics col·lectius del CBUC: antecedents i situació actual". *Item: revista de biblioteconomia i documentació*, n. 41, p. 55-65. <http://www.raco.cat/index.php/Item/article/view/40867/68117>

Benítez, Carles (2006). "Efemèrides del 2006". *Bibliodoc*, pp. 157-177. <http://www.raco.cat/index.php/Bibliodoc/article/download/72734/239098>

Generalitat de Catalunya (2005). "El Dursi impulsa amb 110 mil euros el projecte Revistes Catalanes amb Accés Obert". *Sala de premsa*, 18 marzo. http://premsa.gencat.cat/pres_fs/vp/AppJava/notapremsavw/72171/ca/dursi-impulsa-110-euros-projecte-revistes-catalanes-acces-obert.do

RACO: revistes catalanes amb accés obert. Barcelona: CBUC, 2006. <http://www.raco.cat>

Nota original completa en *Blok de BiD*: <http://www.ub.edu/blokdebid/es/node/668>

* * *

Ciencia abierta: políticas científicas de palos y zanahorias

Tony Hernández-Pérez

*Universidad Carlos III de Madrid,
Departamento de Biblioteconomía y
Documentación
tony@bib.uc3m.es*

Hernández-Pérez, Tony (2016). "Ciencia abierta: políticas científicas de palos y zanahorias". *Anuario ThinkEPI*, v. 10, pp. 260-263.

<http://dx.doi.org/10.3145/thinkepi.2016.57>

Publicado en *Blok de BiD* el 25 de noviembre de 2015



Ciencia abierta

El término "ciencia abierta" fue acuñado en 2003 por el economista Paul David para describir las propiedades de los bienes científicos generados por el sector público, en oposición a la extensión de la percepción de los derechos de propiedad en el área de los bienes

de información.

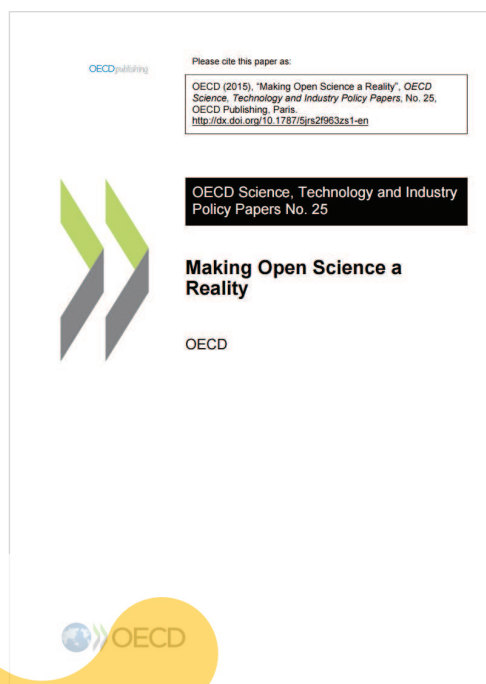
Los economistas consideran el conocimiento científico generado por la investigación financiada con fondos públicos como un bien público, lo que significa que cualquiera debería poder hacer uso de ese conocimiento sin costes adicionales una vez se han hecho públicos, generando así un mayor beneficio social.

El concepto de ciencia abierta se refiere a los esfuerzos que hacen los investigadores, las agencias que financian la investigación y la comunidad científica, incluyendo bibliotecas y centros de datos, para hacer públicamente accesibles en formato digital los resultados de investigación, tanto las publicaciones como los datos de investigación utilizados en la misma, con mínimas o sin restricciones técnicas o legales, lo que permite acelerar las investigaciones en aras de la mejora de la transparencia, la colaboración y el fomento de la ciencia.

Más que artículos y datos

El documento de la OCDE que comentamos en esta nota reconoce también el papel de los ciudadanos, las empresas, las organizaciones sin ánimo de lucro y de los organismos supranacionales en el desarrollo de las políticas de ciencia abierta.

El informe se centra en el acceso abierto a los resultados y a los datos de investigación, presen-



ta las evidencias de los impactos de las políticas científicas para promocionar el acceso abierto y analiza los problemas legales y las soluciones que están adoptando los países de la OCDE para proporcionar un mayor acceso a los datos de investigación.

Se reconoce que la ciencia abierta es algo más que acceso abierto a los datos y resultados de investigación y que incluye también aspectos como la revisión por pares post-publicación, libros de notas de investigación abiertos, software de código abierto, altmétricas, licencias de derechos de autor o ciencia hecha por ciudadanos, como el proyecto *Open Air Laboratories (OPAL)* del Reino Unido, en el que más de 200.000 personas han participado recopilando datos en sus localidades sobre biodiversidad, contaminación de aire o de aguas, etc.

Beneficios

La ciencia abierta no es un fin, es un medio para una ciencia de mejor calidad, más colaborativa y más participativa. Fomentar el acceso a los datos primarios y a los resultados de las investigaciones contribuye a:

- mejorar la eficacia y productividad del sistema científico: reduciendo costes de duplicación en la recogida, creación, transferencia y reutilización de los datos; permitiendo hacer más investigación con los mismos datos; y multiplicando las oportunidades para una mayor participación en los procesos de investigación;
- incrementar la transparencia y calidad de la investigación (permitiendo la replicación y validación de los resultados); acelerar la velocidad

de transferencia (permitiendo mayor innovación);

- asumir retos más globales, como el del cambio climático, que muchas veces requieren amplia infraestructura y coordinación;
- incrementar la conciencia y la implicación ciudadana en la ciencia e investigación (permitiendo mayor conciencia y confianza entre los ciudadanos sobre su verdadero valor).

El informe recuerda que, por ejemplo, el 40% de los usuarios únicos de *PubMed* son ciudadanos, individuos, el 25% procede de las universidades, el 17% de empresas y el resto de organismos gubernamentales y otras categorías.

Políticas

Las medidas que han adoptado los gobiernos para implementar políticas en favor de una mayor ciencia abierta para promover el acceso abierto a resultados y datos de investigación son básicamente de tres tipos:

- regulaciones de mandatos (el palo);
- mecanismos de incentivos (la zanahoria);
- creación de marcos legales e infraestructuras tecnológicas que favorezcan el desarrollo de la ciencia abierta.

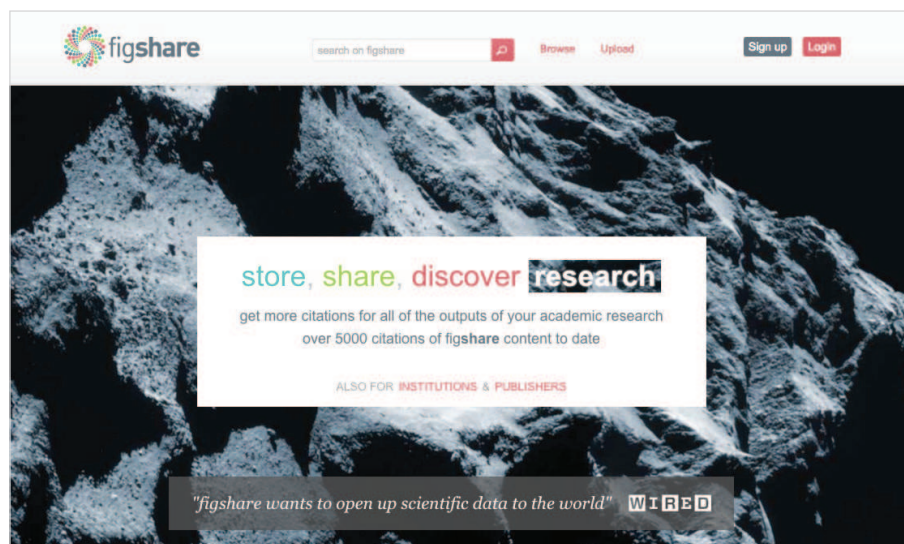
“El concepto de ciencia abierta se refiere a los esfuerzos para hacer públicamente accesibles en formato digital los resultados de investigación, y a las políticas y herramientas utilizadas”

Del análisis del informe se deduce que la mayoría de las iniciativas implementan los mandatos o la creación de infraestructuras y reconoce que se necesitan más políticas de incentivos.

Las políticas de mandatos ya han mostrado sus frutos. Por ejemplo, desde que los *National Institutes of Health (NIH)* implementaron su mandato, la obligación para los investigadores de depositar las publicaciones financiadas con sus fondos, el número de artículos en *PubMed Central* ha aumentado considerablemente: 3,2 millones de artículos en 2014. Y lo más interesante, el número de visitantes únicos diarios se ha duplicado de medio millón a un millón, lo mismo que el de artículos recuperados: en tres años (2011-2014) se ha pasado de un millón a dos millones de artículos recuperados cada día.

Algo parecido ocurrió cuando se liberaron las imágenes del satélite *LandSat* de la NASA. Se pasó de vender 19.000 imágenes al año, que se pagaban a 600 US\$ cada una, a distribuir 2,1 millones

Register for free at <https://www.scipedia.com> to download the version without the watermark



<https://figshare.com>

de imágenes gratuitas al año. Una distribución que genera a empresas como Google Earth un negocio de más de 100 millones al año y repercute positivamente en la economía estadounidense.

Ruta verde y ruta dorada

En cuanto a las publicaciones, los resultados de investigación, los gobiernos, las agencias de financiación e incluso las universidades, están promoviendo tanto los modelos de ruta verde (depósito o archivo en repositorios) como los de ruta dorada (publicaciones gratuitas subvencionadas o con algún sistema de pago por parte del autor). La ruta verde, los repositorios, parece estar quedando como el modelo por defecto para un acceso abierto básico. No obstante, en la mayoría de los países de la OCDE han emergido variantes de la ruta dorada para responder a las preferencias de autores por publicar en revistas líderes que no son, en principio, de acceso abierto y por los intentos de los editores de desarrollar nuevos servicios para hacer competitivos sus modelos de negocio, lo que incluye el apoyo de gobiernos y agencias financiadoras a los modelos de pagar por publicar o *article processing charging* (APC).

Las políticas sobre acceso abierto son muy variadas en Europa, desde la opción de la ruta dorada, por mandato, en los *Research Councils United Kingdom* (RCUK); o por recomendación, en Holanda; a la ruta verde por publicaciones que se impone en Alemania.

Para resolver los problemas relacionados con las leyes de *copyright* en las publicaciones, tanto Alemania como Inglaterra, entre otros países, han hecho enmiendas a sus leyes, básicamente para asegurar, en el caso de Alemania, que el autor de una contribución científica tenga derecho a republicar en acceso abierto el resultado de una investigación si ha sido financiada con fondos públicos después

de un período de embargo máximo de 12 meses. Y ese derecho no puede ser limitado por acuerdos contractuales, incluso si el autor ha cedido sus derechos exclusivos a un editor. En Inglaterra, las reformas a las leyes de *copyright* intentan garantizar una mayor libertad en la reutilización de material científico tanto para educación como para propósitos de investigación no comercial, especialmente para uso de minería de datos y textos.

El informe reconoce que los repositorios y otras plataformas online

de publicación científica no tendrán impacto si la información que contienen no es de buena calidad, si los sistemas de consulta no son amigables y los *datasets* que contienen no han sido debidamente limpiados y curados o los metadatos no han sido suficientemente implementados, por lo que sigue siendo necesaria más formación y más concienciación entre los investigadores para extender la cultura de ciencia abierta.

"La buena calidad de los datos implica no sólo que estén accesibles sino que sean inteligibles, evaluables, fiables y reutilizables"

Datos

Los datos y las mediciones han sido fundamentales para la ciencia y la aparición de nuevos instrumentos y métodos de captura masiva de datos están modificando la forma de hacer ciencia, por lo que el tema de los conjuntos de datos (*datasets*) de investigación tiene ahora incluso más implicaciones para la ciencia. Por ejemplo, el proyecto *Digital Sky Survey*, que comenzó en el año 2000, recopiló más datos con su telescopio en su primera semana que todos los datos que se habían amasado en la historia de la astronomía. O el proyecto *SKA* (*Square Kilometre Array*), cuyo radiotelescopio puede generar hasta 1 *petabyte* de datos cada 20 segundos. Más aún, la capacidad de análisis de datos ha hecho posible que las máquinas de secuenciación de ADN sean capaces de leer 26 billones de caracteres de código genético en segundos.

El informe distingue cuatro tipos de datos de investigación particularmente importantes en investigación:

Register for free at <https://www.scipedia.com> to download the version without the watermark



- Datos observacionales: provienen de telescopios, satélites, redes de sensores, encuestas y otros instrumentos que graban información histórica sobre un fenómeno, incluyendo investigación de ciencias sociales, como encuestas demográficas. Son datos que en muchos casos no pueden ser replicados y deben ser guardados.
- Datos experimentales: los que pueden ser capturados por máquinas de alto rendimiento, como el acelerador de partículas, mediante ensayos clínicos, tests biomédicos, farmacéuticos o experimentos controlados. La preservación de estos datos experimentales es muy importante cuando no es viable o ético replicar esa captura de datos.
- Datos computacionales: generados por simulaciones de ordenadores a gran escala.
- Datos referenciales: datos que son altamente curados y muy demandados por la comunidad científica. Este tipo de datos se crean para diversos propósitos, que van desde el mapeo del genoma humano a las masas de datos longitudinales sobre ciertos aspectos sociales y económicos. El *Worldwide Protein Data Bank* o el *Panel Study of Income Dynamics* serían ejemplos de este tipo de conjuntos de datos (*datasets*) de referencia. Con todos estos datos a menudo es necesario conservar también los materiales auxiliares como la calibración de los instrumentos, los parámetros de los experimentos o los libros de notas.

Más allá de la masiva captura, generación o análisis de datos, el informe aborda el problema de la falta de protocolos estándar de evaluación de calidad de los datos, como existe ya para las publicaciones científicas. Y los datos tienen muy poco valor si no satisfacen un criterio mínimo de calidad. La buena calidad de los datos implica no sólo que estén accesibles sino que sean inteligibles, evaluables, fiables y reutilizables, por lo que es necesario elaborar información sobre compartición de datos y metadatos para un futuro uso de los mismos por múltiples equipos de investigadores.

El problema con los datos de investigación es

que los investigadores no tienen ni las habilidades ni los incentivos que se requieren para llevar a cabo las tareas de curación y difusión de *datasets* ya que se trata de una tarea costosa, que lleva mucho tiempo y que por el momento no conlleva ninguna recompensa para los investigadores. El informe trata dos posibles soluciones al problema mediante incentivos:

- que sean reconocidas las citas de datos (*data citation*)
- que se creen revistas especializadas en publicación de *datasets* o artículos sobre datos (*data journals*).

Y menciona varias organizaciones que están abordando el problema, como *DataCite*, *Orcid*, *Figshare* o *The Dryad Digital Repository*.

<http://www.datacite.org>
<http://www.orcid.org>
<http://www.figshare.com>
<http://www.datadryad.org>

Marcos legales

El informe acaba con un análisis de los marcos legales de la protección de datos en los países de la OCDE, con la descripción de diversas iniciativas sobre ciencia abierta (creación de repositorios nacionales, archivos, centros de datos, plataformas de CVs de investigadores, proyectos de participación ciudadana en la ciencia), tanto gubernamentales (Finlandia, Inglaterra, Canadá, España, Unión Europea, etc.) como de universidades, por ejemplo, el mapa para los datos de investigación publicado por la *League of European Research Universities (LERU)* o por otros centros como los *National Institutes of Health (NIH)* de Estados Unidos.

OECD (2015), "Making open science a reality". *OECD science, technology and industry policy papers*, n. 25, OECD Publishing, Paris.
<http://dx.doi.org/10.1787/5jrs2f963zs1-en>

Nota original completa en *Blok de BiD*:
<http://www.ub.edu/blokdebid/es/node/657>

Register for free at <https://www.scipedia.com> to download the version without the watermark